Vital Annex: International Journal of Novel Research in Advanced Sciences
ISSN: 2751-756X
Volume 03 Number 03 (2024)
https://innosci.org/IJNRAS

Article

Estimating the Number of Students in Najaf Governorate Education Using the LSTM Model and the ARIMA Model

Haitham Hassoon Majid[1] , Hassanien Adel Salih[2]

1,2 General Directorate of Education in Najaf Governorate Najaf, Iraq
* Correspondence: haithamnkhozaee@gmail.com

Abstract: Estimating the number of students is important in educational planning and making the necessary administrative decisions. For these estimates to be accurate, appropriate models must be chosen in the estimation process. The LSTM and ARIMA models are important and commonly applied models in time series forecasting. In this study, these models were used to estimate the number of students in the Najaf Governorate secondary stage, and a comparison was made between these models. In building the ARIMA model, the stability of the series was initially confirmed, and the necessary differences were taken to achieve this stability. Then, a comparison was made between a group of ARIMA models using the BIC, MSE, and AIC criteria to choose the most accurate model in estimating the number of students, and the best model ARIMA (0,2,1) was determined. As for the LSTM model, this model was built using the R program, where the model was trained and evaluated using the study data. The results showed the superiority of the LSTM model in estimating the number of students compared to the ARIMA model.

Keywords: LSTM model, ARIMA models, BIC standard, AIC standard

1. Introduction

The process of estimating the number of students is one of the important educational topics closely related to the academic planning process, and finding the accurate plans necessary to develop the educational process and the process of predicting the number of students is one of the important things required for the process of preparing to provide books and school supplies and providing the teaching staff and other requirements for advancing the educational process. There are many classical methods currently applied to estimate the number of students. Still, these methods may be insufficient and may not give accurate results in many cases compared to modern techniques. To develop educational performance and reach accurate results in estimating the number of students and choosing the appropriate model, a group of important models applied in the analysis of the time series were used in this research, and a comparison was made between these models to choose the appropriate model to predict the number of students (secondary stage) in the education of Najaf Governorate. Two types of models were applied. The first type of model is the Box Genks (ARIMA) model, one of the important and commonly used time series models in predicting various phenomena and analyzing the series. The second type of model is the LSTM model or what is known as the (Long short-term memory) model, which is one of the types of neural networks that can deal with time series data and also can find accurate future predictions, especially in complex series data and long-term series.

(Nam and Schaefer 1995) A study using neural networks to predict passenger traffic was presented, and the study showed the efficiency of this method in estimation.

Citation: Haitham Hassoon Majid. Estimating the Number of Students in Najaf Governorate Education Using the LSTM Model and the ARIMA Model. Vital Annex: International Journal of Novel Research in Advanced Sciences 2024, 3(3), 61-71.

Received: 10th Apr 2024
Revised: 11th Mei 2024
Accepted: 28th Jun 2024
Published: 27th Jul 2024

(Faraway and Catfield, 1998) presented a study on a seasonal time series for an airline using neural networks and Box Jenkins models.

(Kao and Huang, 2000) A study on predicting ambient air quality was presented, in which they used the classical time series method and the neural network method, and the study showed the good performance of neural networks in prediction, even in long periods.

(Baker, 2001) presented an article on identifying factors in educational productivity, in which he used multiple non-linear regression and neural networks, and he explained the importance of this study by adding new features to the model.

(Gamal Al-Shawadfi, 2003) He used hypothetical and real data to present a comparative study on the use of neural networks and the Box Jenkins method in prediction. The study showed the excellent performance of neural networks.

(Al-Otaibi, 2003) used neural networks and Box Jenkins models to predict the series of records in Kuwait and showed that the performance of neural networks is the best in this study.

(Ezza Hazem, Basil Younis 2005) Neural networks were used to predict electricity consumption in the city of Mosul.

(Saliou and Matar 2018) A comparative study on the use of Box Jenkins models and neural networks was presented in the prediction process.

(Ittihad Muhammad Arshan et al. 2020) Box Jenkins models were used to predict the number of students at Ibb University.

(Tian, et al., 2020) She presented a study on predicting COVID-19 infection using a hidden Markov chain model, a hierarchical Bayes model, and an LSTM model. The results showed the excellent performance of the LSTM model compared to the applied models.

(Berat et al., 2021) used the LSTM model to predict the number of students in an American university.

(Jung-Pin et al., 2023) The LSTM model was used to estimate the number of students in primary schools in the United Kingdom. The model showed high efficiency in estimation, with the effect of fertility and migration.

(Hanan Khadari 2023) presented a comparative study between the ARIMA models and the LSTM model to estimate the daily death series due to coronavirus, which showed the superiority of the LSTM model over the ARIMA models in the estimation process.

In this research, the Box Jenkins models will be applied. The appropriate ARIMA model will be chosen for the time series data of the number of secondary school students in Najaf Governorate for the period from 1979-2024, as well as using the LSTM model in analyzing this series and then making a comparison between the proposed ARIMA model and the LSTM model using the BIC, AIC, and MSE criteria.

2. Materials and Methods

The study aims to estimate the number of secondary school students in Najaf Governorate using two time series forecasting models: ARIMA and LSTM. The ARIMA model, a well-known method for time series analysis, involves ensuring the stability of the series through differencing and selecting the best-fit model based on BIC, MSE, and AIC criteria. In this research, the ARIMA (0,2,1) model was identified as the most accurate. The LSTM model, a type of neural network capable of handling complex and long-term dependencies in data, was also employed. The LSTM model was trained and evaluated using Python. Both models were compared to determine their accuracy in forecasting the number of students, with results indicating that the LSTM model outperformed the ARIMA model in terms of prediction accuracy. The study highlights the importance of choosing appropriate forecasting models for educational planning and resource allocation.

3. Results and Discussion

1-1- ARIMA models:

ARIMA models (Auto Regressive Moving Average) are famous and widely applied methods in time series analysis. They are also known as Box-Jenkins models. In this method, the time series is analyzed in some steps before choosing the appropriate model to investigate the time series and make the necessary predictions.

The first of these steps is to draw the time series to ensure that it is stable (time-homogeneous). Although stability is rare, many time series are unstable around the average (there is a certain trend for the series). In this case, appropriate differences are taken to achieve the stability of the series (Box et al., 1976) or that it is unstable in variance. A logarithmic or exponential transformation is performed to achieve stability (Anderson, 1976).

The Box-Jenkins method is one of the important methods in the forecasting process, as it does not impose a specific pattern for the series data before applying the ARIMA model. Still, it starts with an experimental model determined based on the autocorrelation function (A.C.F) and the partial autocorrelation function (P.A.C.F). Then, the parameters of the experimental model are estimated based on the series data, especially those that make the prediction errors as low as possible. Then, some indicators are used to judge whether the model is suitable or inappropriate for the time series. If the model is suitable, we use it to forecast in this series, and if it is unacceptable, the Box-Jenkins method provides information for choosing the appropriate model. The steps will be explained in the next paragraph to clarify this method. Selecting the Box Jenkins Estimate Model:

The first step in determining the model is to ensure whether the time series is stable. If it is stable, we start by choosing the estimated model. However, if the series is unstable, the unstable time series must be transformed into a stable series because stability is the basic condition in applying (ARIMA) models. This is done by taking the appropriate differences (seasonal and non-seasonal) if the series is not stable in the average. However, if the variance is unstable, the appropriate transformations (exponential or logarithmic) are taken for the time series. Sometimes, these differences and transformations are combined to achieve the stability of the time series (Box. G et al., 1976).

After verifying that the series has become stable, the appropriate (ARIMA) model is determined by studying the (A.C.F, P.A.C.F) functions and the appropriate model. There are some important tests applied to verify that the model is suitable for the time series in the Box-Jenkins analysis method, including:

1-2- (Box and Pierce) Test

In this test, a set of autocorrelations of errors is tested once according to the following formula (Shumo, 1987):

$$Q_{BP} = (n - d - SD) \sum_{i=1}^{L} \quad R_i^2(u) \ldots (1)$$

whereas :

n: Number of time series views

d: Number of non-seasonal differences

D: Number of seasonal differences

S: All season

L: Largest displacement

$R_i(u)$: Autocorrelation of errors at displacement (i)

The test hypothesis is:H_0: $R_i(u)= R_1(u)= R_2(u)=\ldots R_L(u)=0$

The test statistic is approximately a chi-square distribution with one degree of freedom. (L-p-q-P-Q)

The decision is to reject the null hypothesis when the calculated statistic value exceeds the table value, and we conclude that the model is unsuitable. Another model is diagnosed and retested, and if the value is less than the table value, then we do not reject the null hypothesis, and the diagnosed model is accepted.

1-3- Ljung-Box Test

This test is used to test any set of autocorrelations in the time series as significantly different from zero. This test can also be used to check the randomness of the data (in general) using a set of test statistic shifts, which can be expressed as follows (Brock Well et al., 1991):

$$Q_{LB} = m(m+2)\sum_{i=1}^{L} \frac{R_i^2(u)}{m-i} \dots (2)$$

Where m=(n-d-SD)

After determining the appropriate model, the parameters of this model are estimated using the maximum likelihood method, the non-linear least squares method, or any other suitable efficient method.

1-4- Some criteria applied to diagnose the ARIMA model:

The test of independence and randomness of the errors resulting from applying the experimental series model is of great importance in determining the appropriate ARIMA model for the series. Still, it is insufficient in some cases, so other criteria have emerged to ensure the model's suitability in analyzing the series and making predictions more accurately. Perhaps the most famous of these criteria are the Bayesian criterion BIC and the Akaike criterion AIC (Brock Well et al., 1991).

1-4-1- Akaike Information Criterion (AIC)

This criterion was proposed by the Japanese scientist (Akaike) in the early seventies to develop the method of analyzing the time series in diagnosing the ARIMA model's rank and the possibility of using it in other areas of statistical analysis. The mathematical formula for this test can be expressed as follows (Akaike, 1974).AIC=m ln(MSE)+V^2……(3)

whereas :

m: Number of views

V: Number of parameters estimated in the model

1-4-2- Bayes' criterion BIC:

Bayes' criterion is a development of Akaike's criterion, where the formula for this criterion can be expressed as follows (Brock Well et al., 1991)

BIC= m ln(MSE)+vln(m)…….(4)

Based on these two criteria, the best model is the model that gives the lowest value in the two criteria.

1-5- Some ARIMA models applied in time series analysis:

These models can only be applied after ensuring the time series is stable. The two functions (A.C.F, P.A.C.F) must be observed according to the model's suitability to the data to identify that the series has become stable.

1-5-1- Generalized Autoregressive Model AR(P):

In this model, the value of the series at time t is expressed in terms of the previous values of the time series plus the value of the random error as follows (Box, G.E. and Price, 1970):

$$Y_t = \mu + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots \alpha_p Y_{p-1} + u_t \dots\dots (5)$$

whereas:

μ : Represents the value of the constant in the model.

α_i : The model parameter represents

u_t : represents the random error of the model such that $(E(u_t) = 0, E(u_t^2) = \sigma^2$, $E(u_t, u_j) = 0)$

The autocorrelation function of this model is exponentially decreasing or sine wave decreasing, and the partial autocorrelation function breaks down after the shift. p (Box, G. E. and Price,1970).

1-5-2- General moving average model (MA(q):

This model's current time series value is expressed using the weighted sum of the prior random errors (Brock Well et al., 1991).

$Y_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q}$ ....(6)

whereas: θ_i Model parameters represent

The autocorrelation function of this model breaks down after the shift q, and the partial autocorrelation function decreases exponentially or sine waves (Wei, W.S., 1990)

1-5-3- Mixed model ARIMA(p, q):

In this model, the two previous models, AR(p) and MA(q), are combined to reach a model with greater flexibility in representing the time series. The general formula for this model can be expressed as (Brock Well et al., 1991):

$Y_t = \mu + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots \alpha_p Y_{p-1} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \cdots + \theta_q u_{t-q}$ ...(7)

This model's autocorrelation function and partial autocorrelation decrease exponentially or sine waves at the shift (p-q) (Wei, W.S.,1990).

1-5-4- LSTM (Long Short Term Memory) model:

Neural networks are considered one of the important and diverse artificial intelligence methods in many fields. They provide a better analytical method for studying the relationship between variables in a more advanced way than traditional methods. Recent computer developments have enabled the application of neural networks in multiple and important fields. Perhaps one of the oldest neural networks that rely on memory is the recurrent neural network (CNN), which appeared in the eighties of the last century. Still, it was limited in application, and the developments witnessed by the world of data and computers in recent years have made CNN technology more important in applied fields. It has the advantage of memory, which enables it to take information from previous data to benefit from it in future predictions. (Lili Mou, Zhi Jin, 2018)

The other type of neural network is the RNN (Rete neural feed-forward) network. This type of network takes a fixed amount of input data at once and produces a fixed amount of outputs as well (Yee 1998), and the LSTM model is considered a type Special from RNN. The LSTM model was first introduced by Hochreiter & Schmidhuber, 1997 as an improvement to the performance of neural networks and to get rid of the problems facing these networks. The LSTM model consists of three parts that work as a neuron carrying the previous state vector. The first part is called the input gate, which adds useful information to the cell. The second part is the forgetting gate, which removes information that has become useless from the cell. The last part is the output gate, which extracts useful information from the cell and presents it as output. (Chimmula and Zhang, 2020) The mathematical formula for these gates at time t can be expressed as follows (Aston Zhang et al., translated by Alaa Taaima, 2022):

$A_t=\alpha(X_t W_{xa}+H_{t-1}W_{ha}+b_a)$....(8)

$B_t=\alpha(X_t W_{xb}+H_{t-1}W_{hb}+b_b)$....(9)

$C_t=\alpha(X_t W_{xc}+H_{t-1}W_{hc}+b_c)$....(10)

whereas:

W_{xc}، W_{xb}،W_{xa} ،W_{hc}،W_{hb}،W_{ha}: Represents the weight parameters

b_a, b_b, b_c: Represents the bias parameters

The LSTM model is important in reducing reliance on long-term information and its negative impact on learning. To design this model, it is necessary to go through five stages: the first is data collection, the second is data processing, the third is data division, the fourth is model improvement, and the last is model evaluation (Omran et al., 2021).

2- The Applied Aspect

In this section, the ARIMA technique and the LSTM technique will be applied to analyze the time series of students' numbers for the secondary stage in the Najaf Governorate Education Department (1979-2014) that were obtained from the records of the Najaf Governorate Education Directorate, and then make future predictions in the time series using the two techniques. Then, a comparison will be made between the results obtained from each method.

The researchers wrote the program in Python to analyze the series using LSTM technology. The SPSS program was also applied to analyze the time series using ARIMA technology.

Finally, the criteria (BIC, AIC, MSE) were used to compare the results of the two techniques.

2-1- Apply ARIMA models:

The first step in the analysis is drawing the time series, as in Figure (1). From the figure, it can be noted that there is a clear trend in the time series with the change in time, which indicates the instability of the series in the average. As for the stability in the variance, from Figure (1), it can be noted that there are no fluctuations in the series that differ from one time period to another. Therefore, the first concept of the time series imposes that the time series is not stable in the average but in the variance.

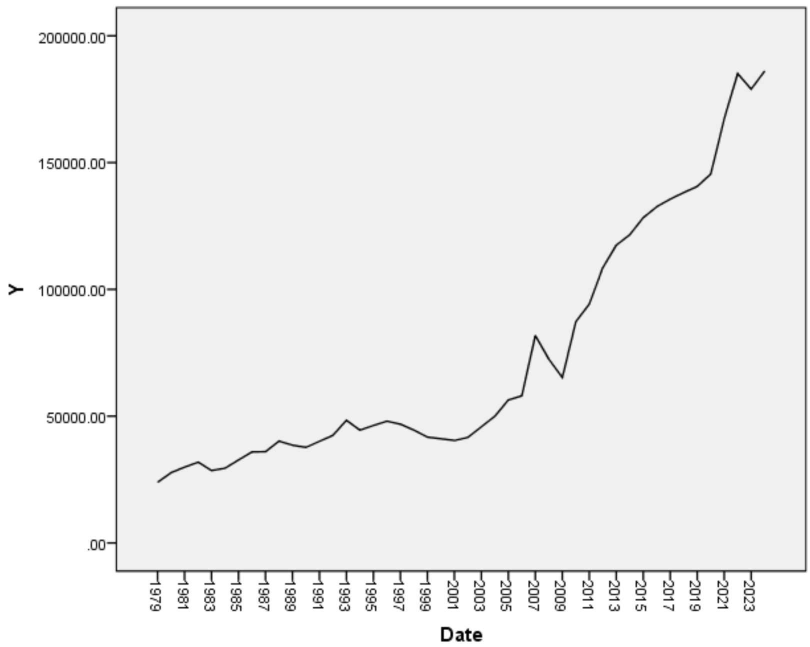

Figure (1) Time series diagram for secondary school students' numbers

The second step in the analysis was to draw the functions (A.C.F, P.A.C) to ensure the stability of the series, as in Figure (2,3) and Table (1), which confirms that the time series is not stable, as the value of the A.C.F, P.A.C.F functions differs from zero even after the sixteenth shift, and the p-value value for the Ljung-Box test is less than 0.05.

Table (1) Autocorrelation function of the series

A.C.F

displacement	Autocorrelation	Box-Ljung Statistic		
		Value	df	Sig. b
1	0.918	41.321	1	0
2	0.839	76.607	2	0
3	0.75	105.479	3	0
4	0.671	129.155	4	0
5	0.609	149.152	5	0
6	0.546	165.601	6	0
7	0.481	178.687	7	0
8	0.414	188.662	8	0
9	0.345	195.771	9	0
10	0.276	200.43	10	0
11	0.206	203.115	11	0
12	0.134	204.284	12	0
13	0.069	204.603	13	0
14	0.024	204.641	14	0
15	-0.019	204.665	15	0
16	-0.045	204.814	16	0

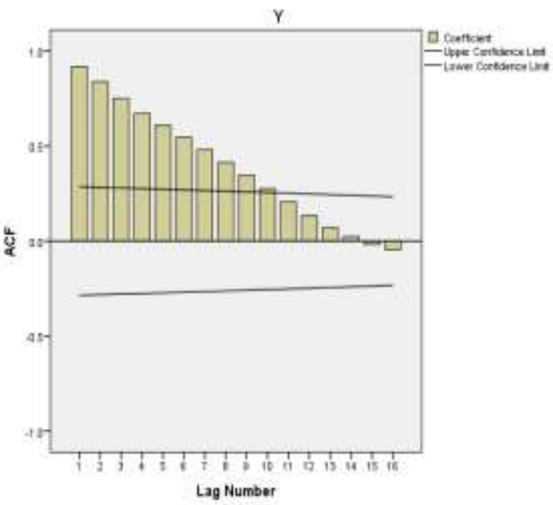

Figure (2) (A.C.F) function

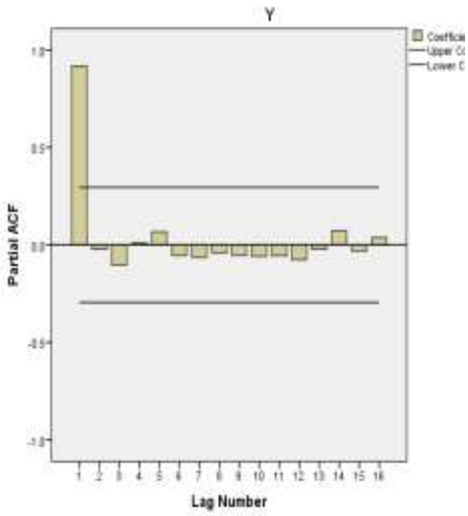

Figure (3) P.A.C.F function

The second difference in the time series was taken to achieve stability. After achieving stability in the series, a comparison was made between a group of ARIMA models to choose the appropriate model for the series, adopting the criteria (BIC, AIC, MSE) as in Table (2), where the ARIMA (0, 2, 1) model was chosen, corresponding to the lowest criteria applied in the comparison.

Table No. (2) shows the comparison between a group of ARIMA models.

The model	**MSE**	**AIC**	**BIC**
ARIMA(1,1,0)	**9838.55**	**30.77**	**31.4**
ARIMA(1,1,1)	**8548.7**	**29.22**	**30.5**
ARIMA(0,1,1)	**8080.4**	**23.9**	**24.3**
ARIMA(0,1,2)	**7822.3**	**19.8**	**20.21**
ARIMA(2,2,0)	**7774.02**	**18.22**	**18.8**
ARIMA(0,2,1)	**6944.44**	**17.34**	**17.94**

After selecting the ARIMA (0,2,1) model, the model parameter was estimated, and its significance was tested, as shown in Table (3). Then, the model was used to predict the number of students for the next five years, and Table (4) shows these values.

Table No. (3) shows the estimation and significance test of the ARIMA(0,2,1) model.

Table (4) shows forecasting student numbers for the next five years using ARIMA(0,2,1)

Table No. (3) shows the estimation and significance test of the ARIMA(0,2,1) model.

	Estimate	SE	p-value
Difference	2		
MA	0.65	0.132	0.004

Table (4) shows forecasting student numbers for the next five years using ARIMA(0,2,1)

	Forecasting the next years using ARIMA(0,2,1)				
	2025	**2026**	**2027**	**2028**	**2029**
Forecast	**196330.5**	**207009.9**	**218175.2**	**229838.22**	**242011**
UCL	**210237.6**	**226899.3**	**242796.3**	**258559.79**	**274437.4**
LCL	**182423.3**	**187120.5**	**193554.1**	**201116.65**	**209584.5**

2-2- Application of LSTM in Time Series Analysis:

The researcher used the R language to write the program. After calling the special libraries in the R program to build the LSTM model and uploading the time series data file, the study data was divided into 80% training data and 20% test data.

The researchers used two layers of the LSTM model, and 128 neural units were in each layer.

The model was trained using the loss function and the optimizer, and then the model was evaluated using the test data and the loss function.

The results showed the ability of the LSTM model to Forecast with high accuracy and with more accurate results than the ARIMA model. Table (5) shows a comparison of the results of the two models in predicting the number of students (within the period)

Then, the forecast was made for the next five years (outside the period) using the LSTM model, as shown in Table (6).

Table No. (5) shows the comparison between the Forecasts using the ARIMA model and the LSTM model

Real values from 1981	Forecast values	SE	Forecast values	SE
	Using ARIMA		Using LSTM	
29886	31464.52	2491725	28213.32	2797858
31886	32802.71	840357.2	30249.22	2679049
28560	34471.52	34946069	32372.92	14538359

29505	29652.93	21883.28	28281.55	1496830
32741	30564.58	4736804	30114.1	6900604
35922	34170.5	3067752	33896.63	4102124
35986	37619.82	2669368	37220.11	1523027
40186	37508.61	7168417	36899.2	10803054
38558	42046.07	12166632	41963.5	11597430
37696	40120.13	5876406	39483.67	3195764
40103	39099.55	1006912	38905.55	1433887
42464	41662.93	641713.1	42027.52	190514.8
48380	44169.24	17730500	44532.55	14802872
44491	50480.66	35876027	51212.78	45182326
46360	46297.82	3866.352	45786.98	328351.9
48045	48287.17	58646.31	48799.64	569481.5
46840	50085.32	10532102	50604.9	14174472
44512	48838.3	18716872	49054.09	20630582
41757	46431.8	21853755	46684.61	24281340
41088	43603.17	6326080	44008.02	8526517
40418	42985.47	6591902	43850.24	11780271
41638	42380.77	551707.3	43332.01	2869670
45823	43761.44	4250030	45029.48	629674
50008	48236.28	3138992	49877.1	17134.81
56405	52707.03	13674982	54214.04	4800306
58048	59472.1	2028061	61143.97	9585030
65263	61299.02	15713137	62120.12	9877695
72504	68903.31	12964968	70446.75	4232278
81800	76527.12	27803263	77844.17	15648591
87241	86268.5	945756.3	87646.06	164073.6
94253	92021.32	4980396	92575.04	2815550
108548	99394.46	83787295	100009.57	72904787
117436	114269.37	10027546	115710.89	2976005
121560	123563.81	4015255	123819.57	5105657
128304	127954.63	122059.4	127274.99	1058862
132595	135043.41	5994712	134622.19	4109499
135614	139614.16	16001280	138642.65	9172721
138185	142884.69	22087086	141594.51	11624758
140608	145702.88	25957802	144240.29	13193531
145469	148377.12	8457162	146789.75	1744381
167262	153556.81	1.88E+08	152222.6	2.26E+08
185137	176077.26	82078889	177083.81	64853869
178934	194574.86	2.45E+08	194435.94	2.4E+08
186125	188438.3	5351357	184237.84	3561373
	MSE=	22175551	MSE=	20658497

Table No. (6) shows the future values predicted using LSTM

	Forecasting the coming years using LSTM				
	2025	2026	2027	2028	2029
Forecast	196967.69	203696.96	212838.78	221019.17	229982.49
UCL	206180.23	218871.51	231341.45	242724.32	254295.23

LCL	187755.15	188522.42	194336.11	199314.01	205669.75

4. Conclusion

The researchers concluded the following: (1) There is an increase in secondary school students annually, as shown by the time series. (2) Using traditional methods to estimate the number of students in the General Directorates of Education, (3) This increase in the number of students requires the development of mechanisms and strategies adopted by the Ministry of Education in developing plans to provide books and school supplies and provide the need for school buildings and other necessities of the educational process. If the estimates are inaccurate, it will lead to a waste of public money.

REFERENCES

[1] A Predictive Model That Aligns Admission Offers with Student Enrollment Probability (Educ. Sci. 2023, 13(5), 440

[2] Akaike , H. (1974) , "A new look at the statistical model Identification " , IEEE Transactions on Automatic Control , Vol.19 , No.6 , PP. 716-723 .

[3] Al-Otaibi, Fawzi: 2003 "Using time series and neural networks in predicting indices" Master's thesis at Al-Ahsa Applied, Suez Canal University, Port Said, Egypt.

[4] Al-Shawadfi, G. A. (2003). A comparison between neural network and Box-Jenkins Forecasting Techniques With Application to Real data. The Research Bulletin, Human Resources Development and Research Center, College of Business and Economics, King Saud University PO. Box.

[5] Anderson , O.D. (1976) , " Time series analysis and forecasting " ,

[6] Aston Zhang et al., translated by Alaa Ta'ima (2022) "Deep Learning" Part Two, Modern Deep Learning Techniques Cairo, Dar Al-Ma'rifah.

[7] Autocorrelations in Autoregressive – Integrated Moving Average Time Series Analysis .

[8] Azza Hazem Zaki, Al-Khayat, Basil Younis Dhnoon (2005), Iraqi Journal of Statistical Sciences, Volume 5, Issue 8, pp. 63-79, 17 pp.

[9] B. Ujkani, D. Minkovska and L. Stoyanova, "A Machine Learning Approach for Predicting Student Enrollment in the University," 2021 International Scientific Conference Electronics (ET), Sozopol, Bulgaria, 2021, pp. 1-4

[10] Baker, B. D. (2001). Can flexible non-linear modeling tell us anything new about educational productivity? Economics of Education Review, 20(1), 81-92.

[11] Box , G. E. and Price , D. A. (1970) , " Distribution of Residual Autocorrelations in Autoregressive – Integrated Moving Average Time Series Models " , JASA , Vol.55 , No.332 , PP.1509-1525 .

[12] Box , G. E. and Price , D. A. (1976) , " Distribution of Residual

[13] Box G. , E. , P. , and Jenkins G. , M. , T. (1976) , " Time Series Analysis" Forecasting and Control " , San Francisco , Holden-Day , U.S.A.

[14] Brock Well , P.J. and Davis , R.A. (1991) , " Time Series Theory and

[15] Butter worths, London and Boston.

[16] Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons & Fractals, 135, 109864.

[17] Faraway, J., & Chatfield, C. (1998). Time series forecasting with neural networks: a comparative study using the air line data. Journal of the Royal Statistical Society: Series C (Applied Statistics), 47(2), 231-250.

[18] Hanan Khadari Mahdi (2023) "Comparison of neural networks with long short-term memory and Box and Jenks models" Scientific Journal of Commercial Research Volume 49 Issue 2 Part One Page 517-542.

[19] Ittihad Muhammad Arshan, Nabil Ahmed Al-Afari, Saba Muhammad Alwan (2020) "Predicting the number of students accepted at Ibb University for the years (2018-2022 AD) using Box-Jenkins models" (Arab Studies in Education and Psychology) Volume 128 Issue 128, Pages 109-129

[20] Lili Mou , Zhi Jin(2018) "Tree-Based Convolutional Neural Networks: Principles and Applications" (SpringerBriefs in Computer Science) 1st ed. 2018 Edition Part of: SpringerBriefs in Computer Science (322 books).

[21] Methods ", 2nd ed , Springer Verlag New York Inc , New York .

[22] Nam, K., & Schaefer, T. (1995). Forecasting international airline passenger traffic using neural networks. The Logistics and Transportation Review, 31(3), 239-252.

[23] Omran, N. F., Abd-el Ghany, S. F., Saleh, H., Ali, A. A., Gumaei, A., & Al-Rakhami, M. (2021). Applying deep learning methods on time-series data for forecasting COVID-19 in Egypt, Kuwait, and Saudi Arabia. Complexity, 2021.

[24] Rahad Imad Saliwa, Dhafer Ramadan Matar (2019) "Comparison of prediction performance between some artificial neural networks and the (Box-Jenkins) methodology with application" Iraqi Journal of Statistical Sciences, Volume 16, Issue 1, Pages 51-76

[25] Shamo, Khaled Dawood ("(1987 Using the time series method to predict the number of monthly deposits at Rafidain Bank", Master's thesis in Statistics, College of Administration and Economics, University of Baghdad.

[26] Tian, Y., Luthra, I., & Zhang, X. (2020). Forecasting COVID-19 cases using Machine Learning models. MedRxiv. 2020 .07.02.20145474

[27] Wei, W. W. S. (1990). Time series analysis: Univariate and multivariate methods. Journal of the American Statistical Association, 33(1), DOI:10.2307/2289741.

[28] Yee, V. P. (1998). Regularized radial basis function networks: theory and applications to probability estimation, classification, and time series prediction. (PhD). McMaster University, , Hamilton, Ontar